DOCUMENT RESUME

ED 161 859 95 SP 013 277

AUTHOR	Lanier, Judith E.; Floden, Robert E.
TITLE	Research and Development Needs for the Advancement of Teacher Education. Research Series No. 8.
INSTITUTION	Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	78
CONTRACT	400-76-0073
NOTE	57p.
AVAILABLE FROM	Institute for Research on Teaching, College of Education, Michigan State University, 252 Erickson Hall, East Lansing, MI 48824 ($3.00)
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	*Curriculum Development; Educational Objectives; *Educational Research; Elementary Secondary Education; Higher Education; Learning Processes; *Professional Education; *Research Utilization; Social Factors; *Teacher Education

ABSTRACT

Three scholarly efforts, in addition to kindergarten through grade 12 research, are necessary before research and development is likely to improve teacher education. First, systematic development efforts must be undertaken to reexamine and translate elementary-secondary research findings into articulate sets of curricula for teachers. Second, research must be conducted to indicate how adults, including, but not limited to, teachers, can be taught the skills found beneficial to their professional activities. Third, systematic attention must be given to the way in which goals for teacher education are determined. The role that each of these three research and development areas must play in improving teacher education programs is examined, and the means by which educational research may be put into practice are discussed. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED161852

Institute for Research on Teaching

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

College of Education — Michigan State University

Research Series No. 8

RESEARCH AND DEVELOPMENT NEEDS FOR THE ADVANCEMENT OF TEACHER EDUCATION

Judith E. Lanier and Robert E. Floden

Published by

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824

February, 1978

This work is sponsored by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Teaching Division of the National Institute of Education, United States Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)

Contents

Abstract iii

Acknowledgements iv

Preface v

Introduction 1

Adapting and Testing the External Validity of Findings from K-12 Research 3

Research on the Education of Professionals 8

Approaches to Research on Professional Education. 14

Research on the Desirable Outcomes of K-12 Teaching and Teacher Education 17

Approaches to the Choice of Desirable Outcomes 24

Behaviors Leading to Student Achievement Gains 28

Judgment Studies 28

Questions in studying human judgment about effective teaching 31

Problems, problems, problems 35

Goals Related to Social Needs 39

Summary and Conclusion 44

Reference Notes 46

References 47

Abstract

Although considerable progress has been made in research on teaching, little advancement has been made in using the results of that research to improve programs of teacher education. The translation of research results into practice can best be expedited by supplementing research on teaching with research and development in three other areas: (1) development work for systematically adapting the knowledge accumulated through K-12 research efforts; (2) research for producing widely applicable knowledge regarding efficacious means of educating professional personnel, particularly teachers; and (3) research and development work for producing tenable specifications of goals for K-12 teaching and teacher education.

The authors discuss the role that each of these three research and development areas must play in improving teacher education programs. The strengths and weaknesses of two different approaches to research in professional education are cited. The two approaches involve (1) descriptive, correlational, and experimental studies following the model of research on teaching, and (2) evaluation studies to examine the joint effects of programs which attempt to incorporate many isolated findings.

Particular attention is given to research and development for producing tenable specifications of goals. The recognized importance of this area is traced in the research literature of the past 30 years. Three approaches to this work are examined: (1) using those teacher behaviors which lead to gains on student achievement measures, (2) accepting the judgments of some group or groups, and (3) determining which outcomes might contribute to the alleviation of specified social needs. The advantages and disadvantages of each approach are discussed, with special attention given to the judgment approach.

Acknowledgements

An earlier version of this paper was discussed at a conference held on October 10-12, 1977, at the Institute for Research on Teaching. We received helpful comments at that time from Gene Glass, Donald Orlosky, Bunnie O. Smith, Richard Turner, and many members of the Institute for Research on Teaching and of the Basic Program Council of the Michigan State School of Teacher Education. Staff members of the Teaching Division at the National Institute of Education not only provided valuable criticisms of our work, but also stimulated our thinking with early drafts of their own papers on research on staff development. Joseph Schwab deserves special credit for directing our attention to the importance of having tenable program goals. Linda Shalaway corrected our English, and Bonnie Klinger provided superlative secretarial assistance. Of course, all credit for errors and omissions must go to the authors.

Preface

Most of this paper should be of interest to everyone concerned with the improvement of teaching; some sections, however, may be more technical than the general reader would prefer. We encourage the reader to move quickly through any section that appears too technical, looking for the major points of the argument without attending to all the details. In particular, we have given detailed attention to judgment studies (pp. 28-39), since they are relatively new to the field of teacher education and readers may desire a more thorough understanding of this research. A careful reading of this section is not necessary, however, for the reader to acquire a sense of our main assertions.

Research and Development Needs for the Advancement of Teacher Education

Judith E. Lanier and Robert E. Floden*

INTRODUCTION

More than 130 years have passed since the first professional school for teachers was established in this country, but prominent educators still report that little is known about how such a professional school should be constituted. Turner (1975) surmised that, "In spite of recent improvements in research in the field, the amount of dependable information available compared to the amount needed to formulate more effective policies and practices in teacher education is miniscule"(p. 107). Many volumes of findings in research on teaching have been published in recent years (e.g., Dunkin & Biddle, 1974; Good, Biddle, & Brophy, 1975; Travers, 1973; Medley, 1977), but the implications of these findings for the improvement of teacher education are unclear. As a result, government officials as well as educational practitioners have become frustrated because the dollars spent for research in education have not produced rapid or sizable improvements in teacher education.

---

*Judith E. Lanier is co-director of the Institute for Research on Teaching, director of the MSU School of Teacher Education, and a professor of curriculum and instruction. Robert E. Floden is a coordinator of IRT's Teacher Education Research Program and an assistant professor of teacher education.

The apparent failure of research to improve practice has prompted a variety of responses, ranging from recommendations for reduced funding of research on teaching at one end, to requests for increased research funding, at the other. The position that support should be decreased is taken by some members of Congress, who insist that if research is to be funded, it ought to improve practice. "As a professional staff member for the Senate Labor, Health, Education, and Welfare Appropriations Subcommittee said last year, 'We want NIE to show us that we are getting the bang for the bucks we are spending on educational research'" (Kerlinger, 1977, p. 8). On the other hand, researchers in teaching argue for increased funding, blaming the general lack of research-based change on the limited amount of research reports. Medley, for example, in his review monograph on Teacher Competence and Teacher Effectiveness (1977) attributes the "weak . . . connection . . . between research in teacher effectiveness and the teacher education curriculum" to "the quality and quantity of research findings to date" and to the "access to these findings" (p. 1).

There is no doubt that both positions have some validity; research on teaching should result in the improvement of teaching practices, and improved amounts and availability of research findings on teaching effectiveness should result in modification of teacher education curricula. Both positions, however, mistakenly place the burden for improving teaching practice on findings from studies of teaching in K-12 classrooms, where the largest and most visible research on teaching has occurred to date.

The authors of this paper posit that such positions are dangerously simplistic and incomplete, leading to unrealistic expectations and inevitable disappointment. Research on teaching effectiveness in K-12 classrooms, alone, can not provide the guidance needed for the improvement of teaching practice. If teaching practice is to be improved, K-12 teaching research must be supplemented by work in at least three additional research and development areas:

1. Development work is needed for systematically adapting the knowledge accumulated through K-12 research efforts, building isolated findings into the larger constellation of expertise teachers need. Such development work should include intensive study of the short- and long-term consequences that the added or deleted knowledge units have on teachers and learners in their natural environment.

2. Research is needed for producing widely applicable knowledge regarding efficacious means of educating professional personnel, particularly teachers.

3. Research and development work is needed for producing tenable specifications of goals for K-12 teaching and teacher education. Such goals necessarily represent judgmental descriptions of what constitutes "success in teaching." It is very likely that different operational definitions of teachers' "work success" have powerful effects on the nature and results of inquiry about teaching.

The activities described above must not be viewed as substitutes for K-12 teaching research, but rather as necessary concomitant inquiries that have significance in their own right, as well as an essential role in the improvement of teaching practice. In this paper we will describe the potential each of these research and development areas has for improvements in teaching. Strengths and weaknesses of alternative approaches in each area will also be noted.

ADAPTING AND TESTING THE EXTERNAL VALIDITY OF FINDINGS FROM K-12 RESEARCH

Historically, teacher educators have drawn on conventional wisdom,

personal experience, common sense, and authority to determine what knowledge and actions were needed for enhancing effectiveness in teaching. There was simply no alternative; sound and useful empirical data were unavailable, and teacher educators had to rely on their own judgment. Recognizing the inadequacy of judgment alone, they naturally lacked self-assurance. Thus, when the first data from K-12 teaching research appeared, the findings were accepted and applied with undue eagerness and haste. Unfortunately, this situation still exists. Findings from K-12 teaching research continue to be accorded greater importance to teacher education than is warranted.

Present results of K-12 teaching studies are still meager and unstable, although more reliable and valid findings are accumulating. No matter how much the methods and results improve, however, these results will not provide knowledge that is directly applicable to the curricular judgments required of teacher educators. Whatever cause and effect linkages might be established between teaching and learning, the findings will generally be limited to maximizing achievement of a single goal. Teacher educators, however, must help teachers cope with competing goals and, in this sense, must be concerned with teaching teachers how to optimize rather than maximize. Unfortunately, the existence of this state of affairs is better understood than is the means to cope with it.

By necessity, any single set of results from research on teaching in K-12 classrooms has limited generalizability. The research process forces a sampling of the complexities of teaching while requiring the researcher to narrowly focus his or her view. The consequence of

narrowing focus is that the researcher must either exclude large segments of potentially important thoughts, actions, and aspects of the teaching environment , or place severe limitations on sample size, so that ethnographic or case study approaches can be employed. In either case, confidence in generalizing findings must be limited.

When findings emerge, then, it is critical that they be applied to, and incorporated into, other settings with existing teaching complexities so that the effects on other teachers, other students, and other relevant variables can be examined. Those variables that were necessarily excluded or ignored in the initial inquiry must now be studied; otherwise, this "obviously good" set of research findings is apt to generate a false sense of confidence.

Take, for example, Rowe's (1974a, b) results from her "wait-time" research. Rowe found that students provided qualitatively better responses to teachers' questions if given more time to think before having to respond. Such results appeared to have immediate relevance for, and transfer to, the curriculum of teacher education. (After all, teachers obviously want their pupils to provide the best possible answers.) Teachers were thus taught to increase the time they provided for pupil responses. But the waiting behavior turned out to be very unstable and short-lived, dropping back when the training and reinforcement periods ended.

While we can only speculate about reasons for the lack of more permanent change in the teachers' waiting behavior, it seems reasonable to assume that some competing goal, such as that of main-

taining group attention, was more important.* The fact that this possibility was overlooked in the application of Rowe's findings is illuminating; it illustrates the need for further examination of the appropriateness of applying K-12 research findings to teacher education curricula.

The response to Flanders' work (1970) provides another example of direct and perhaps premature modification of teacher education curricula on the basis of K-12 research on teaching. Many curricular modifications were made in the content of teacher education to incorporate instruction about Flanders' interaction analysis system. In some cases, for example, instruction on classroom management was discarded and replaced with exercises for learning proper scoring techniques for the Flanders' matrices. Recent time-on-task findings (e.g., Harnischfeger & Wiley, 1976) might raise questions about the appropriateness of this curriculum shift, but these questions aside, it is instructive to consider the assumptions behind this decision to change curricula.

The first assumption was that teachers would be more effective if they were indirect rather than direct, a conclusion that highly overgeneralized the research findings, given the contextual limitations of the empirical work. It was also assumed that adapting and using the research methodology, i.e., the Flanders interaction analysis system, would make teachers more indirect. Finally, it was assumed that

*Various studies of teaching, in fact, (Kounin, 1970; Brophy & Evertson, 1976) have indicated that pupil boredom and lack of involvement leads to distraction and disruptive behavior. Students who wait longer for each other's responses may well become tired with waiting. Thus, while the teacher is helping one student to think and produce better responses, he/she may be inadvertently contributing to the disruptive behavior of others.

whatever was removed from the teacher education curricula to make room for instruction in the Flanders system was less significant to effectiveness in teaching and learning. None of these assumptions was necessarily correct. Systematic development and evaluation would have given decision makers better data and more confidence for exercising their judgment on what to include or exclude in the curriculum.

In addition to validating findings, developmental work would also lead to improved dissemination, i.e., the translation of the findings into more useful and powerful forms. It is common knowledge that verbal dissemination of research findings through written reports or oral presentations at professional meetings is not terribly effective. It is also commonly known that "teaching by telling" and "learning by listening" are typically insufficient means of improving professional practice. Thus, systematic design of instructional strategies, materials, and field trials would help narrow the now classic gap between research and practice.

Feiman (Note 1) emphasized the practical limitations of direct use of K-12 teaching research findings in a comment made during an IRT presentation in 1977: "So Kounin finds that 'withitness' is important for keeping children on task; there's certainly more to the next step than telling teachers that it is important to get with it."

The results of K-12 research on teaching do not directly translate into effective forms of teaching practice. Researchers, practicing teachers, and teacher educators must work together in the mobilization of a well-planned development effort. Such an effort would: (1) accumulate the various findings of K-12 research on teaching; (2) study

the interactions among the variables described in the sets of findings; (3) investigate the degree to which the findings hold in a variety of teaching and learning environments; and (4) incorporate the findings into the larger constellation of knowledge and skills needed by teachers, so that dissemination will be effective. Development work of this kind should facilitate the translation of research into practice.

RESEARCH ON THE EDUCATION OF PROFESSIONALS

Admittedly, research on K-12 teaching will assist in the acquisition of knowledge about effectiveness in teaching children and youths. Emphasis is placed on "assist" to stress the inadequacy of K-12 research contributions, alone, and to imply that because of inherent weaknesses in this body of research, additional modes of inquiry are needed. We emphasize "acquisition of knowledge" to make clear that the availability of knowledge does not alone insure its effect on practice.

We have discussed ways in which development can address the translation and incorporation of K-12 research findings into teacher education. If these steps are taken, the prospects of converting qualitatively better research results into practice should be improved. Again, however, we must stress that the set of necessary inquiries, by itself, will not be enough to adapt sound knowledge into practice.

One area in need of inquiry is research on the teaching of teachers. Knowledge available to the research-on-teaching community should also be available to the K-12 teachers, but we know little about

transmitting the research knowledge in ways that allow teachers to successfully modify their practices. Knowing that the ability to read is a desirable skill to give children does not indicate how children should be taught to read; similarly, knowing that teachers should be taught to provide direct instruction does not indicate how teachers should be taught to provide direct instruction. Without research results on the methods for teaching K-12 teachers, the results of research on teaching cannot be used to improve K-12 teaching. Brophy (Note 2) states the problem succinctly: "Though findings from the research on K-12 teaching effects have important relevance for teacher education, it is unrealistic to seek or expect direct or universal statements about good teaching to emerge from this research."

Some might argue that K-12 research <u>is</u> sufficient to indicate how teacher educators should instruct, and that the research results do have general applicability and can guide the selection of methods of teaching teachers (e.g., some might believe that research showing the value of direct instruction for youngsters is directly relevant for teaching adults). While admitting that K-12 research does not yet provide adequate guidance for teacher education, these challengers might attribute this state of affairs to the size of the research base, implying that what is needed is <u>not</u> a <u>different type</u> of research, but rather <u>more K-12 research</u>.

These arguments can be refuted on two different grounds, the first focusing on correctable weaknesses in current K-12 research and the other centering on weaknesses inherent in K-12 research. The current <u>correctable weaknesses</u> are the limited generalizability across socio-

economic groups and the correlational nature of the results. The inherent weaknesses are the distinctly different subject populations and subject matters of the research, and the influence of teacher training methods on the associations found in K-12 research.

Current research on teaching K-12 is substantially limited to a small segment of the K-12 population.

> For reasons mainly connected with the funding strategy of the U.S. Office of Education, most of the research . . . was done in one segment of the school population -- in classes of Grade III or below in which most of the pupils come from homes of low socioeconomic status. To what extent these findings apply to pupils with other backgrounds or in other grades is not known. What evidence we have about pupils of high socioeconomic status and pupils in the higher grades indicates that results from one group do not always apply to another. (Medley, 1977, p. 2)

Since many teachers are drawn from a higher socioeconomic class than the students studied, results from current K-12 research on teaching might not provide guidance for teaching teachers. However, this lack of generalizability across socioeconomic classes could be alleviated within the existing framework of K-12 research by conducting more research with children from higher socioeconomic groups.

The correlational rather than experimental nature of current research also limits applicability to teacher education. Following the familiar dictum that "correlation is not causation," it cannot be known whether the associations found in the correlational research represent causal relationships or merely some form of spurious correlation. If they are spurious, there is no assurance that changing teacher behavior in particular ways will produce associated changes in children. Although statistical procedures can be used to eliminate some possible sources of spurious correlation, no procedure is foolproof.

Again, however, this weakness in current research could be eliminated within the framework of K-12 research by conducting experimental studies. In fact, investigators have recently begun to conduct experimental tests of the associations discovered in K-12 research (Borg, 1970; Gage, 1976; Gage & Crawford, Note 3). Such investigations are frequently linked closely with developmental efforts such as those described in the previous section.

Although the two difficulties in K-12 research identified above can seemingly be alleviated by modifications within the K-12 framework, other obstacles to application cannot be overcome without employing another type of research. These obstacles are: (1) the problem that results might not generalize to an older population learning a different subject matter, and (2) the problem that all methods of teaching particular behaviors or skills may not necessarily reap the previously-associated advantages.

It is difficult to believe that results obtained with students in elementary school will apply equally well (if at all) to adults of college age preparing to teach (preservice) and to older adults already teaching (inservice). One of the major lessons derived from recent developmental psychology is that children are not merely little adults. Children have their own ways of thinking and learning; introspection about adult perceptions cannot inform us about the perceptions of children. A corollary here is that adults are not merely big children. Adults have different perceptions and needs than do children, and a strategy successful for teaching children will probably not be successful for teaching adults. Medley (1977) pointed out that re-

sults obtained from young elementary school children probably do not apply to students in the higher grades. That assertion is even more relevant when the students are adults. No amount of research in K-12 can determine whether the results of K-12 research will be applicable to older groups.

Differences between subject matter taught in K-12 and in programs of teacher education also limit the applicability of K-12 findings to preservice and inservice teacher education. The standard curriculum in elementary and secondary school does not include subject matter such as classroom management, educational psychology, or methods of teaching mathematics. These subject matters bear little resemblance to any part of the K-12 curriculum, and no amount of K-12 research can discern the direction or extent of the differences; hence, a different type of research is necessary to provide information on how teachers can best be taught.

Finally, the instructional framework in which teachers acquire behavioral modes may alter the effects their behavior has on the children they teach. Return, for a moment, to the example of Rowe's work. The effect of wait-time on school children may be substantially different if teachers are taught not only the importance of wait-time, but <u>also</u> to consider when waiting is most appropriate (as opposed to being behaviorally conditioned to wait three seconds for a response to a question). The reflective teacher may occasionally realize that his/her question was poorly phrased and rephrase it immediately. The behaviorally-trained teacher would just wait three seconds, in which time a student might make a digressive response. Divergent K-12 re-

search results induced by different methods of educating teachers cannot be investigated within the framework of K-12 research; the investigation of the differing effects requires a move into research on the professional education of teachers.

Research which would avoid the difficulties inherent in K-12 research need not be restricted to teacher education. While K-12 research cannot be generalized to teacher education because of the differences in age and subject matter, supplemental research in areas of teaching with similar age groups and subject matters might well be applicable to teacher education. Areas of professional education such as medical education, the education of clinical psychologists, and teacher education may be sufficiently similar for results in research on teaching in one area to be applicable to the other areas. In each of these areas, adults are taught some combination of discipline knowledge and technical skills in a setting which combines forms of didactic group instruction and supervised field experience. Although the areas undoubtedly differ in other respects, these striking similarities in age level, subject matter, and instructional settings make a high degree of generalizability across areas likely, or at least much more likely than generalizability from K-12 teaching to teacher education.

By broadening the scope of research to include all these areas of professional education, investigators could profit from many new research findings and from the variation in specific instructional techniques that other fields have developed. Assuming that the findings in research on other areas of professional education are relevant to and perhaps even equally valid for teacher education, progress in teacher education could be expedited. In addition, if instructional techniques

from the other areas were introduced into teacher education, the benefits to teacher education could be increased manyfold.

Why, it might be asked, should this extension be limited to research on professional education? Why not carry it further, to encompass all research on adult learning and on the teaching of adults? After all, the age groups are the same in each case.

The answer here is obvious. Although the age groups are similar, the subject matters and instructional settings vary widely. Because of these substantial differences, results would probably not be generalizable between professional education and more general adult education. Restricting the scope of research to professional education strikes a balance between similarity in situations and broadness of situation. The scope would be narrow enough to avoid a breakdown in generalizability, yet broad enough to enable researchers to find audiences for their work and to find a sufficient variety of meaningful problems.

In conclusion, we believe that the current weak link between K-12 research and the practice of teaching and teacher education would be strengthened by research on the education of teachers. It is our further contention that this research might be conducted most expeditiously by including in its scope all types of professional education.

Approaches to Research on Professional Education

Two general approaches may be taken toward research on professional education, each with distinct strengths and weaknesses. The first approach follows the general methods of research on teaching K-12 learners and attempts to investigate the separate effects of distinct

components of teaching through descriptive, correlational, and experimental studies. The second approach attempts to assess the aggregate effects of a complex combination of components through program evaluation. The research approaches complement each other, and both can contribute to the improvement of teacher education.

The approach taken in K-12 research has the advantages of an analytical approach. Components are considered separately, and the effects of an entire program can often be predicted from knowledge of those components, provided that interaction effects are relatively small. This approach has as an added asset, familiarity. Many behavioral scientists understand the techniques involved and, as a result, might be able to apply them to the area of professional education.

Research on teaching K-12 has often been severely criticized, however, particularly on the grounds that studies have been almost exclusively correlational, rather than experimental.* The lack of experimental studies may be due to the great difficulty in obtaining experimental control over the assignment of treatments to students. Although this control was achieved in the Follow Through Study (Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977), it was accomplished only at great expense. Control over the assignment of subjects may be easier to obtain in research on professional education, where scheduling is less difficult and the cooperation of the participants and appropriate administrators is more easily obtained.

---

*This criticism may subside as more experimental studies are conducted. It is interesting to note that the most prominent experimental studies that have been or are being conducted -- Borg, Kelley, Langer, & Gall, 1970; Gage, 1976; Gage & Crawford, Note 3 -- would be best classified as research on professional education. The only experimental intervention lies in the teaching of professionals. That is, there is no direct intervention in the classroom; changes are to occur only indirectly as a result of the teacher education.

If interactions among components of an educational experience are powerful (as suggested by Cronbach, 1975), studies of only a single factor have limited value in determining the effects of that factor in the context of a complex educational program. More precisely, research which considers only main effects has limited predictive power when interactions are large. Information about interactions can be obtained by investigating the joint effects of components in a more elaborately designed study. However, the number of subjects needed for a study of even a small number of factors and their interactions quickly becomes prohibitive.

Program evaluations can be used to assess the combined effects of a complex combination of components, although the individual contributions of these components cannot be determined. The Follow Through Study might better be thought of as a program evaluation than as a research study investigating individual factors; the same might be said of the current study by Gage and Crawford (Note 3).

Program evaluations also provide information on the components as actually implemented, rather than as experimental manipulations in a research setting. Often, components change in subtle (and not so subtle) ways when they are administered as part of a larger program by people who are administrators rather than researchers.

The disadvantage of program evaluations is that they yield results which often have extremely limited generalizability. The evaluation might indicate how well an entire program worked in its particular setting, but it has little power to predict how parts of the program would work when adapted to other settings. Still, program evaluations are essential

to the improvement of teacher education, since every program of professional education is composed of many components within an administrative structure; it is unreasonable to believe that the interactions, which are likely to be large, will have been assessed by previous research.

Again, it is important to emphasize the promise of, and need for, coordinated efforts in this area. If common metrics are agreed upon for evaluating various teacher education programs, and if data are accumulated and shared, they can readily be studied with new meta-analysis techniques (see Glass, in press). New approaches would afford increased potential for generalizing findings from program evaluation studies.

RESEARCH ON THE DESIRABLE OUTCOMES OF K-12 TEACHING AND TEACHER EDUCATION

Our third and final suggestion for building more effective knowledge bases in teaching -- perhaps the most difficult and challenging one -- is to develop tenable definitions of "effective teaching." While this task is probably the most critical one for improving research in teaching both K-12 and professional populations, it is the one most neglected by the research and development community. No research on K-12 teaching or study of professional teacher education (preservice or inservice) can ever be used to construct better programs for teaching teachers unless some decision is made about the definition of "better," and teacher education programs cannot be compared in quality until such decisions are made explicit.

Partial definitions of what constitutes effective teaching can be

inferred from the goals teacher education programs seek to accomplish. Those goals, in turn, should determine the outcomes measured when program effects are studied. Justification of the particular choice of outcomes is seldom made, but outcomes must be chosen before a program is evaluated or research conducted. We argue that: (1) the choice of outcomes implies selection of definitions of effectiveness, and (2) the choice should be explicit and carefully justified, i.e., the definitions should be tenable.

Researchers on teaching focus on teaching outcomes and characteristics that, in some sense, they consider important or desirable. The outcome variables or salient characteristics selected by researchers are either de facto operational definitions of work success in teaching, or they are expected to be related to work success.

Operational definitions chosen by the investigator for purposes of scientific inquiry may or may not be valued by others in the educational community. For example, standardized tests are often used as outcomes measures in research on K-12 teaching, because they are thought to measure student attainment on generally recognized educational goals. There is, however, a notable lack of consensus among educators on the adequacy of standardized tests as measures of teaching work success. Nevertheless, a number of investigators focusing on K-12 research continue to rely heavily on these tests.

No standardized tests are available to measure desirable outcomes of teacher education. The National Teachers Examination does not have a broad base of support, so researchers and designers in teacher education typically look elsewhere for guidance. They frequently rely on "expert" judgments of supervisors and administrators. In some cases,

researchers investigating the teaching of teachers avoid the question of what constitutes success in teaching and select a "proxy" variable or set of variables, instead. Often, these proxy variables simply show relationships among elements in the training program (e.g., success in one course related to success in an ensuing course). In essence, they substitute success in the program, itself, for work success in teaching. In such cases, attention is given to program components, and the choice of teaching success outcomes is left either to the judgments of researchers on K-12 teaching or to whoever supervises prospective teachers during the practice teaching experience.

The little research on teaching inservice teachers that has been conducted has also tended to neglect the development of tenable definitions of teaching work success. There are, of course, some notable exceptions, such as the work of Borg and his colleagues (Borg, Langer, & Wilson, 1975; Borg, 1976).

The major emphasis in research on teacher education, then, has been on teacher preparation and on describing and analyzing program elements of particular practical interest. Thus, most teacher education research has ignored the problem of choosing tenable outcomes, concentrating instead on procedural questions such as: (1) the amount of subject matter or educational foundations knowledge needed for success in a particular program, (2) the effects of diverse types and amounts of field experience on student and supervisor judgment, (3) the effects of competency-based teaching methods, (4) modeling and feedback procedures, and (5) simulated training exercises on student achievement of course objectives.

It is important to note that the focus tends to be on methods of teacher preparation. Although research into such methods has obvious importance for decisions about instructional programs, it is highly debatable whether useful information can be gathered without a more tenable specification of program goals. Questions on methods of teacher preparation are relevant and, as yet, unanswered, but the answers will tell us little about what facilitates effective teaching until the goals of teacher education have been more carefully determined.

Turner, in his "overview of Research in Teacher Education" (1975), supported this view. He stated: "The objective of analyzing teacher work success is to clarify what one is trying to optimize" (p. 87). He also supported a shift in research emphasis from the focus on program components to the outcomes.

> Moreover, the classical view of research in teacher education has not shown remarkable results over the past 30 years, and an alternative to it may be regarded as desirable from the prospective of the overall research strategy in the field, the optimization of teacher work success. (p. 91)

Turner was by no means the first to recognize the urgent need for clarity of goals in research on teaching. A generation earlier, Rabinowitz and Travers (1953) had stated:

> Any study of ability depends upon a conception of what constitutes successful functioning. Before definitive research on the factors associated with the effectiveness can be pursued, it must be possible to specify some criterion through which effectiveness may be identified. Research based upon a clearly unacceptable criterion cannot produce results of any great significance. Why then do we not develop more adequate criteria of teacher effectiveness? (p. 212)

Six years before that, Flanagan (1947) noted:

> It is impossible to study the requirements for success in

> an activity without defining the activity. A complete definition of what is meant by success in the activity is practically identical, with a statement of the procedure for obtaining a criterion.

As recently as 1976, at the meeting of the Invisible College of Researchers on Teaching (held at the Institute for Research on Teaching), prominent members of the research community agreed that a major reason for disappointing research results was the weakness in conceptual clarity about outcomes and criterion measures.

Of central interest here is that the neglect of outcomes in research on teaching is not due to a lack of awareness of the problem, or even to disagreement (at a general level) about the form of the solution. Why, then, has the recommendation been acknowledged so strongly in words, yet ignored in deed?

Thirteen years after Rabinowitz and Travers (1953) urged researchers studying teaching to begin concerning themselves with criterion questions and necessary value judgments, Travers reflected on the lack of progress toward this end. In his most salient observation, he urged researchers to avoid predictably "dead-end" approaches to outcome specification, such as the standard behavioral objectives approach and the Bloom-Krathwohl type of taxonomic work. Commenting on the weakness of the behavioral objectives approach, Travers (1966) stated:

> Since Tyler many years ago wrote on the importance of defining educational objectives in behavioral terms, vast lists of behavioral outcomes have been prepared. Those who have engaged in this activity seem to agree on the virtues of undertaking this task and also on the unwieldly and unuseable product which generally emerged from it. (p. 113)

It should be noted that this warning was issued prior to the emergence

of the competency performance-based movement in teacher education. Yet, awareness of the problems and probable outcomes, alone, was insufficient to prevent the production of seemingly endless lists of disorganized bits of "desirable" outcome behavior.

Travers (1966) also condemned taxonomies built around "response inferred" characteristics. He criticized the selection and organization of tasks on the basis of response-inferred characteristics (such as knowledge and comprehension) just as strongly as he admonished the collectors of behavioral objectives.

> They are of little use in the development of theory of instruction. Since the taxonomies classify tasks on a basis which lacks any established scientific utility, the application of the taxonomies results in a bulky and unmanageable collection of tasks which generally cannot be used for scientific purpose. The product which results is reminiscent of those bags of unsorted foreign postage stamps which philatelists like to rummage through in the hope of finding some rare item. (p. 114)

To avoid the behavioral objective problem, Travers recommended the construction of a taxonomy of tasks with a scientific base and scientific utility for building theories of teaching. He suggested the use of a "system of scales" -- a set of high-order instructional tasks selected on the basis of a value judgment of their importance to learning specific content. He urged researchers to give special attention to subject matter specificity, as opposed to taking the generic approach of many competency-based efforts. By this suggested approach, all lower-order tasks would be subsumed beneath the higher-order ones; the scaling would define the outcomes, and the resulting ordered systems of tasks would, in turn, define the outcomes of instruction.

Travers (1966) described another way of viewing the process:

> All outcomes of education can be viewed as the consequences of the production of rule-regulated behavior. Outcomes can be specified in terms of the rules that regulate behavior. These rules can represent narrow rules which specify the orderliness of behavior with respect to single very narrow tasks, in which case the outcomes of any educational program have to be specified in terms of a large number of rules. On the other hand, the rules may be broad in scope and refer to general lawfulnesses covering a wide range of tasks -- in which case the objectives can be specified in terms of a set of relatively few rules. While procedures for defining objectives in the past have implied that many rules have to be specified, the proposal here is that these be organized into broader rules covering a wider range of events. (p. 115)

A review of the professional activity that followed Travers' recommendation reveals no attempts to act upon his advice; instead, researchers have engaged in a cycle of activities: (1) reasserting that clarity of goals is important, (2) getting others to agree, (3) making lists of important outcomes, (4) reviewing the results and deciding that the effort was a waste of time, and (5) concluding that the original idea to clarify goals was a bad one.

The conclusion, however, should not be that the concept of clarifying outcomes is inappropriate, but rather that the attempts at implementation have been faulty. We agree with the scholars of each decade who have urged pursuit of more appropriate definitions of work success in teaching. Further, to avoid repitition of the mistakes of the past, we call for an indepth analysis of several alternative approaches to goal specialization, with the hope that they will provoke both criticism and creative suggestions. The problem is not an easy one, but we are optimistic; the task we face, although complex, does not appear impossible, and we are hopeful that new technologies may help

us find the clarity we need.

Approaches to the Choice of Desirable Outcomes

Three general methods of choosing desirable outcomes may be identified: (1) using those teacher behaviors which lead to gains on achievement measures, (2) accepting the judgments of some group or groups, and (3) determining which outcomes might lead to the fulfillment of specified social needs. Each of these approaches -- along with its advantages and disadvantages -- will be discussed.

Behaviors Leading to Student Achievement Gains

Outcomes for teacher education are most commonly selected from those teacher actions and behaviors which have been shown to produce gains on pencil and paper achievement measures, particularly on standardized achievement tests in reading and mathematics. Examples of this approach can be found in almost all articles reviewing research on teaching in K-12 (Medley, 1977; Rosenshine & Furst, 1973; Rosenshine, 1971, 1976) and in many texts on educational psychology (Gage & Berliner, 1976; Good & Brophy, 1977). Not all researchers, of course, consider standardized tests the best indices of desirable outcomes. Kounin (1970) and Borg (1976), for example, are more concerned with teaching actions which decrease student disruptive behavior; such studies are relatively rare, however. To simplify this exposition, we will concentrate only on standardized tests of reading and mathematics as guides for choosing outcomes.

The assumption behind this method of selection is that gains on

reading and mathematics tests are desirable educational outcomes. While these gains are not necessarily the only desirable outcomes, they are the most easily measured and have become the primary ones toward which teaching actions have been geared. Proponents of this approach admit that the behaviors associated with these gains will not produce all the desirable educational outcomes; but they argue that at least some desirable outcomes will be produced and challenge others to propose a method which can do more. This approach may not lead to the best possible teacher education program, but it seems likely to produce improvement and allows for immediate action. Results that have been obtained in K-12 research can be used immediately to establish outcome measures for research on professional education.

This approach to choosing outcomes, however, can be challenged on at least three different points. First, are the tests used reasonable measures of the basic skills that have values generally agreed upon? Second, does a generally-valued set of basic skills even exist? Third, is a general agreement that the basic skills are valuable a sufficient reason to believe that they are, indeed, valuable?

Standardized tests have received substantial criticism in recent years on the ground that they are poor measures of the valued basic skills.

> The standardized achievement tests are also highly correlated with standardized intelligence tests, thus causing us to wonder exactly what kinds of items are really used in these tests. Furthermore, the tests are usually group administered multiple-choice tests. When working with young, bilingual, or lower socioeconomic status children, there is a serious question about whether many of the children are being appropriately tested. (Berliner, 1975, p. 4)

In The Myth of Measureability (Houts, 1977), a collection of articles

taken from the Principal, several authors suggest that close examination of standardized tests reveals the tests bear little relation to skills and concepts valued by the public. While these claims might be somewhat exaggerated, they do indicate that standardized tests should not be blindly accepted as appropriate measures of education.

A related, but even deeper criticism of this approach to choosing outcomes is that there is little agreement about which particular set of skills should be considered basic. In mathematics, for example, considerable attention was given during the 1960s to the question of whether or not material such as elementary set theory was part of the mathematical basic skills. The question of the composition of the basic mathematical skills was never answered, however. As recently as October, 1975, a National Institute of Education Conference directed itself to the question, "What are basic mathematical skills and learning?" (NIE, 1975).

> It is proposed here that the determination of what mathematics is most worth learning is a task that will require careful and systematic study from the perspectives of several interest groups. (Helms & Graeber, 1975, p. 70)

> The challenge to describe basic skills and learning in school mathematics is an assignment full of pitfalls. In the past five years, hundreds of mathematics educators, school systems, professional groups and the National Assessment have been busily composing taxonomies of fundamental objectives for mathematics instruction at various grade levels. With few exceptions, these efforts to establish a reasonable list of basic skills have been failures. There has been no general agreement among the competing groups. Moreover, the implementation of the various lists as curriculum guidelines threatens to produce fragmented mathematics programs that resemble occupational training more than they resemble education in mathematical methods and understandings likely to be of long range value. (Fey, 1975, p. 51, emphasis added)

These mathematics educators may have exaggerated the differences of opinion concerning composition of the basic skills; however, their comments do suggest that no particular set of skills has general sanction.

A further weakness with the use of standardized tests as guides for choosing outcomes is that the different tests do not even measure the same skills. Recent analyses indicate that the most widely used tests cover significantly different content (Porter, Schmidt, Floden, & Freeman, 1978). Hence, even if some single set of basic skills could be identified, not all the major achievement tests would measure progress on that set of skills.

Finally, it must be recognized that general popular sanction is not necessarily an adequate reason for accepting an outcome as valuable. The fact that a large number of people might esteem an outcome does not automatically confer value upon that outcome. Here, "valuable" is taken to mean "beneficial to members of society as specified by explicitly stated criteria," rather than "capable of being valued."

The distinction between "valued" and "valuable," or "desired" and "desirable," or "wanted" and "worthwhile" becomes clearer if one considers that the majority of people may often be led to value particular outcomes by advertising and publicity. Such fabricated values are unstable and probably do not reflect the real beliefs of the people.

Popular opinion is not the determinant of valuability, yet the argument for using tests because they are accepted is based on the assumption that popular opinion <u>does</u> lead to value. Critics of the accepted-test approach question the assumed connection between valued

and valuable and insist that some alternative reason must be found for choosing the outcomes and goals of teacher education.

In summary, identifying goals for teacher education by association with student gains on standardized tests has the advantages of familiarity, convenience, and general acceptability. But this approach can be attacked on the ground that the outcomes have not been adequately defended as valuable, i.e., the outcomes are not tenable. A similar trade-off will be seen in the discussion of the next general approach for determining the appropriate outcomes of teacher education.

Judgment Studies

A second approach to determining appropriate outcomes of teacher education is to identify and accept the best judgments of some group or groups of individuals deemed worthy of choosing outcomes. Such groups might include parents, teachers, school officials, teacher educators, behavioral scientists, or some other constituency. The judgment of a particular group might be obtained by direct questioning or indirectly through a questionnaire or behavioral study.

This approach has two obvious strengths. First, it automatically generates political support for the program of research and development which follows. If the goals for teacher education are solicited from a certain group, that group would seem bound to support research and programs striving toward those goals. Second, the judgment-studies method resolves one of the major shortcomings of the standardized test approach. By empirically determining what goals are valued, the question, "Do people really value the goal?" is answered in the affirmative (at least as far as the empirical procedures are successful).

Furthermore, since administrators and teacher educators must make judgments about goals in their current work, anyway, those day-to-day judgments should at least be made explicit. By casting these judgments into a framework of empirical study, greater clarity can be obtained about who the judges are and what basis they are using for judgment.

For purposes of teaching teachers or studying the teaching of teachers, an explicit conception of teaching that presents the accompanying "ought" views appears critical. Rabinowitz and Travers (1953) recognized this 25 years ago and cited many experts before them who shared their view:

> There is no way to discover the characteristics which distinguish effective and ineffective teachers unless one has made or is prepared to make a value judgment. The effective teacher does not exist pure and serene, available for scientific scrutiny, but is instead a fiction in the minds of men. No teacher is more effective than another except as someone so decides and designates. Teachers are real enough, and methods are available or can be improvised to study these real teachers. But the effective teacher is only an abstraction. The process of designating any particular teaching practice as effective or ineffective inevitably stems from a reasoned judgment. The ultimate definition of the effective teacher does not involve discovery but decree . . . In the final analysis a criterion is based not upon evidence but judgment. But let the experts speak:
>
> > "In the development of an original criterion . . . validity is a logical rather than a statistical concept." (Bechtoldt, 1947, p. 359)
> >
> > "The criterion . . . can be subjected to no wholly satisfactory empirical test of its adequacy. The criterion must, consequently, be logically justifiable as valid in its own right." (Brogden & Taylor, 1950, p. 160)
> >
> > ". . . there is no way in which criteria of success can be established on an empirical basis. The definition of success in any activity must always be based

> on rational considerations." (Flanagan, 1947, p. 162)
>
> "The ultimate criterion of success in any duty must always be determined on rational grounds. There is no other basis on which this choice can be made." (Thorndike, 1949, p. 123)
>
> "The particular statements of what constitutes a good teacher in any particular locality are in the nature of policy statements -- emphasizing those qualities which are deemed to be acceptable to the person or group whose thinking has dominant force. . ." (Scates, 1950, p. 141)

If this reasoning is correct, then our first step should be directed toward scientific studies of human judgment. Though the technology for such efforts is still in its infancy, new and improved methods show increasing promise. Hammond and Adelman (1976), in "Science, Values and Human Judgment," discuss how the integration of facts and values requires the scientific study of human judgment:

> Current efforts to integrate scientific information and social values in the forming of public policy are confused and defeated by the widespread use of ascientific methods -- the adversary system and the person-oriented approach. The adversary system suffers from an ascientific commitment to victory rather than truth; the person-oriented approach suffers from an ascientific focus on persons and their motives rather than on the adequacy of methods. The reason for the widespread use of both lies in the failure to recognize that human judgment can be brought under scientific, rather than ad hominem, analysis. The argument advanced here is that a scientifically, socially, and ethically defensible means for integrating science and human values can be achieved. (p. 395)

The comments of Hammond and his colleagues illustrate a developing technology and growing enthusiasm for scientific inquiry in areas dealing with the integration of scientific fact and human values. The psychological study of judgment may provide the technology needed to describe the policies of important educational constituencies. Because no such technology is yet fully developed, however, the tools

that are available must be closely examined before they can be used with confidence.

One framework within which the empirical study of policies could be conducted is a system of scales, as described by Travers (1966). The system would allow individuals who value different outcomes to locate points of agreement and clarify differences,but would not itself prescribe outcomes of primary importance. Again, a study of the individuals to be served by the research would have to be undertaken to determine the particular policy or policies to guide the choice of outcome variables.

Researchers often use their own judgment and policies for choosing outcomes, although those judgments and policies are seldom formally described. While this procedure may be appropriate at times, it makes the comparison and cumulation of findings exceedingly difficult. It would be helpful, it seems, if researchers would make their own judgments and policies explicit and indicate how they compare with the policies of other groups, such as teacher educators, elementary and secondary school teachers, parents, and government officials. Research using outcome variables suggested by these alternative policies might find a larger appreciative audience.

Questions in studying human judgment about effective teaching. Do teacher educators or parents have identifiable policies? Many people assume that they do, with minor individual differences evident between groups. Disputes between or among representatives of the various groups are often explained in terms of these perceived differences in position , e.g., "Teacher educators believe that teachers should be

expert in the subject matter areas, while teachers feel the need for greater emphasis on classroom management." Such statements assume that the differences within a group -- among teacher educators, for example -- are smaller than the differences between groups -- between teachers and teacher educators, for example. Unfortunately, there is little evidence to support that assumption. Similarly, although many people feel they can identify the policies of different groups, their descriptions of those policies are based only on personal experience and the literature of professional groups. The choice of outcome variables in a broad research plan should have a firmer foundation.

The techniques that Hammond and Adelman (1976) used in their psychological research on judgment were directed at assessing individual policies about type of bullets that should be used by the local police. Hammond determined the relative weight attached to factors of central importance, considered those weights to assess each of the bullets in question, and suggested a decision satisfactory to all parties. Although the elements in Hammond's study are more clear-cut than are the corresponding elements in the planning of research on teaching teachers, his work suggests a potentially productive approach.

A naive method of determining people's judgment-making policies is to ask them directly what factors they consider in making judgments. Parents, for example, might be asked how they decide which teacher they prefer for their child, with the researchers probing to reveal the factors (or "cues") considered and their relative weights. A list of

possible cues might be presented to the parent to provide additional structure to the procedure. In either case, policies of several parents might be aggregated to a single "parent judgment policy."

This approach has three serious flaws. First, if the interviews are unstructured, aggregation may be very difficult; it would probably be difficult to extract specific cues from a lengthy narrative and even more difficult to compare the cues extracted across narratives. In addition, people are notorious for making exaggerated claims about the complexity of their decision processes. Finally, it is well known that people do not make accurate predictions of their own future behavior. In real life, a parent might choose a teacher quite different from the type indicated to the researchers.

To avoid the drawbacks of soliciting direct descriptions of judgment, psychologists have adopted a more sophisticated method of determining judgment-making policies. They ask research subjects to make judgments based on descriptions of complex situations. Statistical procedures are then used to infer the relative importance of the factors used by the subjects in making the judgments. For example, parents might be asked to rate teachers based on a series of descriptions. If the researcher constructed the descriptions by intentionally varying a set of cues, he might use regression analysis to describe relationships between the cues and the judgments. Such a study is called a policy-capturing study.

The policy-capturing approach provides at least a partial solution to the first two difficulties of the direct approach. Responses are easily classified since the cues under investigation are determined in advance. More importantly, a better assessment of the com-

plexity of judgment is obtained because the judgments, rather than the subjects' description of the judgments are used as data.

The third difficulty -- the discrepancy between judgments made in research settings and "real world" judgments -- may be partially overcome by making the research setting as similar to the "real world" as possible. Brunswik and his followers have advocated use of this procedure, emphasizing the losses entailed by any other approach. In addition to making the judgment situations as life-like as possible, they choose only combinations of cues which are encountered outside the research setting. (A competing school of research holds that the advantages of independently chosen cues outweigh the disadvantages posed by occasional strange cue combinations.)

Several recent studies have dealt with teacher judgments (e.g., Shavelson, Cadwell, & Izu, 1977), and one has directly addressed judgments of teacher quality (Anderson, 1977). Anderson's study provides a good example of the general approach to be attempted, but it focuses on the general process of judgment rather than on the factors contributing to particular judgments. The report centers on factors influencing consistency of judgment (e.g., verbal vs. numerical presentation of cues describing the teachers to be judged). In marked contrast, a study being considered by the Teacher Education Research Program at the Institute for Research on Teaching would probe the selection of factors used in making judgments and compare the patterns of judgment among important groups in education.

As a look at the past failures reveals, the task of adequately analyzing ends is extremely difficult. Although the introduction of

new technologies provides some hope for eventual success, new technologies also introduce new difficulties. Some difficulties can be anticipated and should be considered.

Problems, problems, problems. Judgment studies have only recently been applied to the identification of value positions. A number of questions still need attention before the approach can be applied with confidence to the choice of goals for teacher education. We will cite several of these problem areas (without providing solutions) to indicate the scope of work which must be undertaken before this approach can be used.

First of all, completeness and precision of expression are desirable features of a system of scales, but features that are difficult to obtain. An ideal system would include all features that constitute judgments of teachers. A scale constructed only on the basis of abstracts from the literature can suffer from an overly narrow view. Although the system can always be adjusted to incorporate new material, research planners must stop reading at some point and start planning. How will they know when that point is reached?

The system should also provide precise characterizations of each of its component scales so that interpretation of the scales will be uniform. To understand the difficulty of this task, one need only consider the volumes of philosophical literature attempting to clarify concepts such as "liberal education" or "the inquiry approach." When we consider the difficulty philosophers have had in making one idea precise, we become aware of the pressing need for precise devices

to simplify characterizations of the major dimensions along which teachers are judged.

The successful construction of a system of scales is still important, however, for the policy-capturing approach. Yet, one of the greatest difficulties in conducting a policy-capturing study is determining which cues to investigate. We have proposed that the system of scales be used as a source of cues; in that case, the development of a precise system must precede the choice of cues. If the system is inadequate, what other sources of cues should be used?

Even when the system is satisfactorily constructed, it will undoubtedly suggest far more cues than any single policy-capturing study can investigate. If only 10 cues were considered (each of which could be either present or absent), 1,024 different possible descriptions could be generated. Since many scales would give a wider range of options, it is clear that only a small fraction of the system could be investigated at one time. How should the choice of cues for initial investigation be made, and how should additional scales be incorporated into subsequent investigations?

Construction of the appropriate judgment task is also a difficult undertaking. It should be as realistic as possible, yet flexible enough to be appropriate for several different groups of individuals. The precise wording of the judgment question will undoubtedly have a large influence on responses received. In preliminary work at the Institute for Research on Teaching, the need for precision and clarity was repeatedly emphasized. How should the judgment task be constructed and what should that crucial question be?

Once the policy-capturing study is successfully completed, a

larger and more nebulous problem must be resolved. The results of the policy-capturing study must be used to select outcome variables in planning research on teaching teachers; the identification of value positions alone is insufficient for identifying outcomes. How will the link between "captured" policies and outcome variables be forged?

Concomitant with the preceding problem is the question of measurement. As the entire area of affective measurement testifies, important outcomes cannot always be easily measured; some people, in fact, suggest that only trivial outcomes can be measured precisely. The problem of measurement has been avoided in this discussion, but must be faced if practical research is to be planned. The question of whether or not attention should be limited to currently measurable outcomes must be answered.

Components of judgments may be determined by methods other than direct and indirect interview. In an ongoing study, for example, Florio (Note 4) is using ethnographic techniques to identify the salient characteristics of the performance of a teacher highly regarded in the community. While addressing the question of what aspects of the teaching appear salient, Florio does not try to elicit the information directly from parents or other groups, as in the policy-capturing studies previously described. Rather, she attempts to describe the thoughts and actions of a teacher whose work is known to be valued by particular groups.

A quite different approach to choosing goals through the study of

judgments is to use favorable judgments, themselves, as goals of teacher education. As is commonly done in research on teacher education, one can design a program which would identify teachers who are highly valued by other professionals or by students. Much of the research reviewed by Turner (1975) uses these unanalyzed judgments as the outcome variable. While the policy-capturing and ethnographic approaches attempt to break down judgments of teachers into the components which are desirable or undesirable, the use of judgments as outcomes makes no such attempt at analysis. Favorable judgments by supervisors, administrators, or students are seen as desirable ends in and of themselves and, therefore, no analysis is necessary. In such cases, however, what the judges believe, or perceive, to constitute "effective teaching" is left unknown.

In addition to the methodological problems encountered in obtaining accurate reports of judgments, the judgment study approach is subject to two of the major criticisms of the standardized test approach. First, the measures used in research or program development may not accurately reflect the goals after the goals are determined. One reason why basic skills receive so much attention in educational research is that it is relatively easy to measure the extent to which they are learned. Yet, as we know, measures of achievement in the basic skills are soundly criticized; one can only imagine how much stronger the criticism would be of measures developed to assess progress directed toward other goals.

Second, and perhaps more importantly, the judgment study approach, like the standardized test approach, seems to rest on the dubious

assumption that popularly-supported outcomes are valuable. As we have explained, valued outcomes are not necessarily valuable. Since this distinction and its implications were previously discussed in depth, they will only be briefly sketched here. Basically, the criticism rests on the belief that most people have not given sufficient attention or research to educational goals to be able to make good judgments about goals. The judgments they make are probably diluted versions of the judgments of influential educators and politicians. (A better approach might be to examine the statements of these influential individuals.) In any case, an appeal to the wisdom of the masses is not a particularly good approach to the selection of goals for teacher education.

In summary, the judgment-studies approach has the advantage of generating political support and popular approval. It also establishes a procedure for choosing goals which can be completed in a relatively short time period. It can be criticized, however, on the grounds that it is not likely to lead to goals which can be practically measured and it fails to provide valid reasons for the choice of goals. The final approach examined will have quite a different set of strengths and weaknesses.

Goals Related to Social Needs

The final means of selecting appropriate goals of teacher education is to (1) determine pressing societal and community needs, (2) determine which needs might best be met through the educational system, and then (3) choose teacher education goals which will help meet those needs. The major advantage of this approach is that it

provides an adequate justification for the goals chosen. The major problem is that the approach is time-consuming and difficult. In Teachers Make a Difference, Good, Biddle and Brophy (1975) describe the complicated and uniquely American (U.S.) context from which this problem emerged:

> Once we had little difficulty stating the goals of education . . . The teacher's task was to train pupils in the time-honored Three R's, and teachers who failed in this mission . . . were deemed by their constituents as failures.
>
> Today, the problem is no longer as simple. For one thing, success in the modern world requires more than simple literacy. For another, the school has grown in size and complexity, and it is tied to other schools as part of a complex system of education. As a result, today we demand that the school accomplish a broader range of more difficult tasks. This would be challenge enough, but we are also less certain about what tasks should be assigned to the school. Society is ethnically diverse and rapidly changing, so that we cannot be sure that the attitudes and skills we demand of pupils today will serve them twenty years from now . . . Thus, not only is the catalogue of goals for the school broader and more complex, there is less consensus concerning these goals within society.
>
> Nor does the problem stop with the fact that we hold confusing goals for our schools. Worse yet is the fact that we have few clear mechanisms for resolving conflicts among goals, and often we lack sufficient empirical information to know what to do even if we could all agree on established goals. Most other Western societies have well-established mechanisms to debate and make decisions about educational goals. Unfortunately, our mechanisms for this purpose are weak. Moreover, as education becomes more complex, as our goals for it become more ambitious, educators and other citizens become more confused about how to accomplish these goals. Questions of this sort require research, and too often the needed research is not available. (emphasis added, p. 87)

Indeed, it seems that if practice and research in teaching and teacher education are to be significantly advanced, continuing and major efforts must be made to provide improved knowledge in this area.

There are a number of strengths and weaknesses in choosing goals related to social needs. Some of them can be readily seen by considering how the approach might be carried out.

Social and individual needs are, in every case, the reference points to which any adequate justification of an educational objective must be attached. Learning to add, for example, is not, in itself, valuable; learning to add is valuable only insofar as it helps meet an individual need (such as balancing a checkbook) or a social need (such as facilitating the transfer of goods). It is not adequate to say that a goal is valuable because some people happen to value it; it <u>is</u> adequate to say that a goal is valuable because it helps satisfy a social or individual need. Hence, the determination of educational goals in general, and goals of teacher education in particular, should end with a reference to individual or social needs.

A possible first step in selecting goals, then, would be to identify social and individual needs. Since there is often a period of years between the time formal schooling ends and the time when needs are realized, the needs should either be constant or "on the rise" in order to be relevant when the child has grown. Such needs might be identified by asking prominent economists, sociologists, psychologists, and educators to apply their knowledge of recent American history to the task of predicting which social and individual needs they think will be most pressing in 15 years.

There are several other possible ways of selecting goals which would satisfy needs. One would be to begin with an analysis of the present goals of the school and the societal needs they are apparently

trying to address. Another way might be to conduct a set of historical inquiries that would seek to identify needs which have been an enduring part of our society. Both approaches would concentrate on our uniquely American situation because we are chiefly concerned (at least in this paper) with education in the United States.*

Following the sociological functionalists (esp. Parsons, 1951), Schwab (Note 5) has suggested the following rubrics for organizing social needs; he claims that society needs to:

1. Be reasonably coherent and conflict-resolving.
2. a. Impart a sense of membership in some group.
   b. Impart a sense of membership in a collection of subgroups we call American society.
3. a. Produce useful social roles for most of its members.
   b. Have most of the useful social roles reasonably well played.
4. Have a reasonable number of its members leading reasonably satisfactory lives.
5. Have a reasonable quantity of expertise.
6. Have a quantity of invention and deviation to insure instigation of dissatisfaction with the status quo.

These rubrics are, of course, too broad to serve as the basis for determining educational goals. The categories are only meant to suggest needs which might then lead to educational goals.

A further factor comes into play when attempts are made to determine the goals of education and teacher education through study of societal needs. One must consider the relative efficiency and

*One might argue that the United States is too diverse a society for such a study and that the scope of application ought to be restricted. The trade-off between general applicability and congruence to the specific needs must be considered.

appropriateness of schooling as compared to alternative social institutions in meeting the needs identified.

The task set by the social needs approach is indeed large. It would, perhaps, be unmanageable if no goals could be identified and used until the entire task was completed. The practical solution is to proceed one need or goal at a time. When a single pressing need is identified, associated educational goals may be established, and alternative ways in which teacher education can be used to meet the goal can be investigated. Incremental rather than radical changes can be made in the teacher education curriculum.

The obvious strength of this approach is that it is designed to provide -- and does provide -- an adequate answer to the question, "Why is this educational goal valuable?" The answer will always be: "Because goal G leads to social or community need N being met, and schooling is a relatively efficacious way to meet this need." Further, the social needs approach would subject the goals of education, the purposes of American schools, and the expected outcomes of school teaching to continuous critical examination for their realism of scope and for their relevance to the changing needs of American society.

The major problems with this approach are that it is time-consuming and difficult to carry out. Unlike other approaches, this one calls for a continuing (in the sense of never-ending) set of inquiries. Even the initial efforts would require a substantial and indeterminate period of time for identifying various procedures and methods of goal selection and specification; the identification of needs is difficult because an adequate definition of "need" is lacking and because virtually any acceptable definition would be hard to put into

practice. It is hard to determine the needs (as opposed to wishes or desires) of an individual, and even harder to determine the needs of a group.

In summary, the needs approach to the identification of goals answers the crucial "why" question, but a price is paid both in magnitude and in difficulty of task.

SUMMARY AND CONCLUSION

The apparent failure of research efforts to significantly improve teacher education programs has been blamed on an insufficient quantity of good research and on the failure of teacher educators to use established results. The solutions often suggested are more money for research on teaching in K-12 classrooms, and improved dissemination of the findings to teacher educators.

A different analysis is suggested in this paper. We suggest that three scholarly efforts, in addition to K-12 research, are necessary before research and development is likely to improve teacher education. First, systematic development efforts must be undertaken to reexamine and translate the K-12 research findings into articulate sets of curricula for teachers. Second, research must be conducted to indicate how adults -- including, but not limited to, teachers -- can be taught the skills found beneficial to their professional activities; we have called this research on professional education. Third, careful and systematic attention must be given to the way in which goals for teacher education are determined. We call this determination of teacher education outcomes. The lack of previous work in these three research and development areas provides an alternative explanation for

the lack of research-based teacher education. Our suggested solution to the problem is to seek support and initiate inquiry in these three domains.

For each of the necessary research and development activities, we have explained the unique contribution the effort would make to improvements in teaching practice. Further, we have provided several different ways in which work in each area could be conducted. All suggested approaches have notable strengths and weaknesses, with no alternative considered uniformly better than the others. This suggests that work should be conducted on each of the approaches suggested. The weaknesses of one alternative will be at least partially overcome by the strengths of the other approaches, and better balance can be maintained. Some of the approaches will have immediate implications for teacher education, while others will take considerably longer to come to fruition. Work in some areas has already begun; work in others has not yet even been planned.

Until further work is conducted on (1) the systematic development of K-12 research findings into curricula for teachers, (2) research on professional education, and (3) determination of teacher education outcomes, no amount of research on K-12 teaching will significantly advance teacher education. The research-to-practice gap is no longer simply a function of inadequate findings in K-12 teaching research or poorly distributed publications. The most "bang for the bucks" will now be produced by concomitant work in the three suggested areas.

Reference Notes

1. Feiman, S. Unpublished presentation, Michigan State University, 1977.

2. Brophy, J. Personal communication, Michigan State University, 1977.

3. Gage, N., & Crawford, J. Study in progress at the Stanford Center for Educational Research, Stanford, California, 1977.

4. Florio, S. Work in progress at the Institute for Research on Teaching, Michigan State University, 1977.

5. Schwab, J. Presentation at Teacher Education Research Conference, Institute for Research on Teaching, Michigan State University, October, 1977.

References

Anderson, B.L. Differences in teachers' judgment policies for varying numbers of verbal and numerical cues. Organizational Behavior and Human Performance, 1977, 19, 68-88.

Bechtoldt, H.P. Problems in establishing criterion measures. In D.B. Stuit (Ed.), Personnel research and test development in the Bureau of Naval Personnel. Princeton, N.J.: Princeton University Press, 1947.

Berliner, D. Impediments to the study of teacher effectiveness (Tech. Rep. No. 75-11-3). San Francisco, California: Far West Laboratory for Educational Research and Development, 1975.

Borg, W.R., Kelley, M.L., Langer, P., & Gall, M. The mini course: A microteaching approach to teacher education. Beverly Hills, California: MacMillan, 1970.

Borg, W.R., Langer, P., & Wilson, J. Teacher classroom management skills and pupil behavior. Journal of Experimental Education, 1975, 44, 52-58.

Borg, W.R. Changing teacher and pupil performance with protocol. Journal of Experimental Education, 1976, 45, 9-18.

Brogden, H.E., & Taylor, E.K. The theory and classification of criterion bias. Educational and Psychological Measurement, 1950, 10, 159-186.

Brophy, J., & Evertson, C. Learning from teaching: A developmental perspective. Boston: Allyn and Bacon, 1976.

Cronbach, L.J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Dunkin, M., & Biddle, B. The study of teaching. New York: Holt, Rinehart & Winston, 1974.

Fey, J.T. Remarks on basic skills and learning in mathematics. In The NIE Conference on Basic Mathematical Skills and Learning, Vol. I: Contributed position papers. Washington, D.C.: National Institute of Education, 1975.

Flanagan, J.C. Personnel psychology. In Current trends in psychology. Pittsburgh: University of Pittsburgh Press, 1947.

Flanders, N.A. Analyzing teaching behavior. New York: Addison-Wesley Co., 1970.

Gage, N. A factorially designed experiment on teacher structuring, soliciting, and reacting. Journal of Teacher Education, 1976, 27, 35-38.

Gage, N., & Berliner, D. Educational psychology. Chicago: Rand McNally, 1976.

Glass, G.V. Integrating findings: The meta-analysis of research. Review of Research in Education (Vol. 5), in press.

Good, T., Biddle, B., & Brophy, J. Teachers make a difference. New York: Holt, Rinehart & Winston, 1975.

Good, T., & Brophy, J. Educational psychology: A realistic approach. New York: Holt, Rinehart & Winston, 1977.

Hammond, K.R., & Adelman, L. Science, values and human judgment. Science, 1976, 194, 389-396.

Harnischfeger, A., & Wiley, D.E. The teaching-learning process in elementary schools: A synoptic view. Curriculum Inquiry, 1976, 6, 5-43.

Helms, D., & Graeber, A. Problems related to children's acquisition of basic skills and learning of mathematics and some suggested R&D options for NIE support. In The NIE Conference on Basic Mathematical Skills and Learning, Vol. I: Contributed position papers. Washington, D.C.: NIE, 1975.

Houts, P.L. (Ed.). The myth of measurability. New York: Hart Publishing Company, 1977.

Kerlinger, F.N. The influence of research on education practice. Educational Researcher, 1977, 6, 5-12.

Kounin, J. Discipline and group management in classrooms. New York: Holt, Rinehart and Winston, 1970.

Medley, D. Teacher competence and teacher effectiveness. Washington, D.C.: AACTE, 1977.

Miller, D.M., & Lutz, M.V. Item design for an inventory of teaching practices and learning situations. Journal of Educational Measurement, 1966, 3, 53-61.

National Institute of Education. The NIE Conference on Basic Mathematical Skills and Learning, Vol. I: Contributed position papers. Washington, D.C.: NIE, 1975.

Parsons, T. The social system. Glencoe, Illinois: Free Press, 1951.

Porter, A.C., Schmidt, W.H., Floden, R.E., & Freeman, D.J. Impact on what?. The importance of content covered (Res. Ser. No. 2). East Lansing, Michigan: The Institute for Research on Teaching, Michigan State University, 1978. Paper presented at the 1977 meeting of the Evaluation Research Society, Washington, D.C., October, 1977.

Rabinowitz, W., & Travers, R.M.W. Problems of defining and assessing teacher effectiveness. Educational Theory, 1953, 3, 212-219. Reprinted in R.T. Hyman (Ed.), Contemporary thought on teaching. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Rosenshine, B. Teaching behaviors and student achievement. London, England: National Foundation for Educational Research in England and Wales, 1971.

Rosenshine, B. Classroom instruction. In N.L. Gage (Ed.), The psychology of teaching methods: The 75th Yearbook of the National Society for the Study of Education (Part I). Chicago: University of Chicago Press, 1976.

Rosenshine, B., & Furst, N.F. The use of direct observation to study teaching. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Rowe, M.B. Wait-time and rewards as instructional variables, their influence on language, logic and fate control. Part I: Wait-time. Journal of Research on Science Teaching, 1974, 11, 81-94. (a)

Rowe, M.B. Wait-time and rewards as instructional variables, their influence on language, logic and fate control. Part II: Rewards. Journal of Research on Science Teaching, 1974, 11, 291-308. (b)

Scates, D.E. The good teacher establishing criteria for identification. Journal of Teacher Education, 1950, 1, 137-141.

Shavelson, R.J., Cadwell, J., & Izu, T. Teachers' sensitivity to the reliability of information in making pedagogical decisions. American Educational Research Journal, 1977, 14, 83-98.

Stebbins, L.B., St. Pierre, R.G., Proper, E.C., Anderson, R.B., & Cerva, T.R. Education as experimentation: A Planned variation model. Volume IV-A: An evaluation of Follow Through (Rep. No. 76-196A). Cambridge, Massachusetts: Abt Associates Inc., 1977.

Thorndike, R.L. Personnel selection: Test and measurement techniques. New York: John Wiley & Sons, Inc., 1949.

Travers, R.M.W. Towards taking the fun out of building a theory of instruction. Teachers College Record, 1966, 68, 49-60. Reprinted in R.T. Hyman (Ed.), Contemporary thought on teaching. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Travers, R.M.W. Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Turner, Richard L. An overview of research in teacher education. In K. Ryan (Ed.), Teacher education: The 74th Yearbook of the National Society for the Study of Education (part 2). Chicago: University of Chicago Press, 1975.